Part 1

Визуальная Одометрия – это процесс оценки движения камеры в среде с помощью средств передвижения (транспортное ср-во, человек или робот) использующий в качестве входных данных одну или несколько камер, прикрепленных к одному их этих средств. Прикладные области включают робототехнику, носимые вычислю устройства, дополненную реальность, автомобили. Термин Визуальная Одометрия был поляризован в 2004 году Нистером в его революционной статье [1], но уже были изданы более ранние работы [88], [89].

Этот термин был выбран из-за его схожести с методом Колесо Одометрии, который постепенно оценивает движение транспортного ср-ва с помощью встроенного счетчика вращений этого колеса в течении времени. Аналогичным образом, метод Визуальной Одометрии управляется постепенной оценкой положения транспортного средства через проверку изменений, которые производятся движением на картинках от бортовых камер. Для эффективной работы метода Визуальной Одометрии должна быть достаточная освещенность окружающей среды и статичная сцена с достаточными текстурами, чтобы очевидное действие могло быть определено. Более того, последовательность кадров должна быть зафиксирована с обеспечением того, что они имеют достаточное перекрытие сцены. Преимущество метода Визуальной Одометрии, с уважением к методу Колесо Одометрии, в том, что метод Визуальной Одометрии не подвержен влиянию неровной местности или других неблагоприятных условий.

Это демонстрирует, что по сравнению с методом Колеса Одометрии, метод Визуальной Одометрии обеспечивает более точную оценку траектории, с вероятностью ошибки в местоположении от 0.1 до 2%. Это возможность сделать метод Визуальной Одометрии интересным дополнением к методу Колеса Одометрии и, дополнительно, другие навигационные системы, такие как глобальная система местоположений(GPS), внутренняя система измерений (IMUs), и метод Лазерной Одометрии (похож на метод Визуальной Одометрии, Лазерная Одометрия оценивает движение камеры транспортного средства внутри среды путем сопоставления последовательных лазерных сканов). При отказе GPS в окружающих средах, таких как подводная или воздушная, Визуальная Одометрия имеет крайнюю важность.

Две части этого пособия и их обзор обеспечат широкое введение в метод Визуальной Одометрии и исследование, которое проводилось с 1980 по 2011 год. Хотя первые два десятилетия доказывают множество независимых реализаций, только во время третьего десятилетия сделали преуспевающую систему, работающую в реальном времени, что привело к тому, что метод Визуальной Одометрии будет использоваться на других планетах двумя Марсоходами-вездеходами впервые. Первая часть этого пособия повествует об истории первых 30 лет исследований в этой области и об ее основах. После короткого обсуждения моделирования и калибровки камеры, она описывает основную оценку (pipeline – потока данных??) двумя схемами Монокулярная и Бинокулярная[[1]](#footnote-2), описание плюсов и минусов каждой реализации. Вторая часть о функциях соответствия (сцен, картинок), надежности и применения. В нем будут рассмотрены распознавание характерных точек, используемых в методе Визуальной Одометрии, и алгоритм Фильтрации (Outline-Rejection[[2]](#footnote-3)). Особое внимание будет уделено обсуждению методу *RANdom SAmple Consensus[[3]](#footnote-4)* (RANSAC) и различным модификациям для его ускорения. Другие затронутые темы будут отображать ошибки, распознавание местоположения (или обнаружение замыкания цикла) и узел регулировка.

Это учебное пособие написано как для опытных, так и для начинающих пользователей с рекомендациями и ссылками на алгоритмы для создания полной системы VO. Поскольку идеального и уникального решение VO для любых рабочих сред не существует, оптимальное решение должно быть выбрано тщательно в соответствии с конкретной навигационной средой и данными вычислительными ресурсами.

**История метода Визуальной Одметрии**.

Проблема восстановления относительной позы камеры и трехмерная (3-D) структура из набора изображений камер (калиброванные или некалиброванные) известна в сообществе компьютерного зрения как структура движения[[4]](#footnote-5) (SFM). Его происхождение можно отнести к работам, таким как [2] и [3]. VO - частный случай SFM. SFM более общий и решает проблему трехмерной реконструкции как структуры, так и положения камеры представляют из последовательности упорядоченного или неупорядоченного набора изображений. Окончательная структура и положение камеры обычно уточняются с помощью автономной оптимизации (т. е. корректировка узла), чье время вычисления растет с количеством изображений [4]. И наоборот, VO фокусируется на оценке трехмерного движения камеры последовательно - по мере поступления нового кадра - и в реальном времени. Регулировка узла может быть использована для уточнения локальной оценки траектории.

Проблема оценки движения транспортного средства только из визуального ввода появилась в начале 1980-х годов и была описана Моравеком [5]. Интересно заметить, что большинство ранних исследований в VO [5] - [9] было выполнено для планетарных вездеходов и мотивировало программу NASA по исследованию Марса в стремлении предоставить вездеходу всю местность с возможностью измерять ее в 6-ти степенях свободы (DoF – Degree of Freedom) при наличии проскальзывания колеса на неровных или грубых местах.

Работа Моравека выделяется не только представлением первой оценки движения с помощью потока кадров, чьи основные функциональные блоки все еще используются сегодня, но также описанием одного из ранних угловых детекторов (после, впервые предложенный в 1974 году Ханной [10]), который известен сегодня как угловой детектор Моравека [11], предшественник, однажды предложенного Форстнером [12] и Харрисом, и Стивенсом [3], [82].

Моравек тестировал свои работы на планетарном вездеходе оборудованным тем, что он назвал стереослайдером: одна камера, скользящая по рельсу. Робот двигался в стиле «stop-and-go», оцифровывая и анализируя изображения в каждом положении. На каждой остановке, камера скользит горизонтально, снимая девять снимков в равноудаленные интервалы. Уголки были обнаружены в изображении используя его оператор и согласованный вдоль эпиполярных линий остальные восемь кадров с использованием нормализованной взаимной корреляции.

Потенциальные совпадения местоположения у следующего робота были снова найдены корреляцией, используя стратегию «coarse-to-fine» подхода, чтобы учесть масштабные изменения. Затем выпадающие результаты были впоследствии удалены с помощью проверки глубинных несоответствий в восьми стереопарах.

Наконец, движение было рассчитано как неподвижное тело для выравнивания триангулированных трехмерных точек, наблюдаемых на двух последовательных позициях робота. Система уравнений была решена с помощью взвешенного наименьшего квадрата, где веса были обратно пропорциональны расстоянию от трехмерной точки.

Хотя Моравек использовал одну скользящую камеру, его работа принадлежит классу стереофонических алгоритмов VO. Эта терминология объясняет тот факт, что относительное трехмерное положение признаков непосредственно измеряется триангуляцией в каждом месте робота и используется для получения относительного движения. «Trinocular» (Тринокулярные) методы относятся к одному и тому же классу алгоритмов. Альтернативой стереозрению является использование одной камеры. В этом случае доступна только информация об опоре. Недостатком является то, что движение можно восстановить только до масштабного коэффициента. Затем абсолютную шкалу можно определить из прямых измерений (например, измерения размера элемента в сцене), ограничений движения или интеграции с другими датчиками, такими как датчики IMU, давления воздуха и дальности. Интерес к «monocular» (монокулярным) методам объясняется наблюдением того, что стерео VO может выродиться в монокулярный случай, когда расстояние до сцены намного больше, чем «stereo baseline» стереофоническая линия (т. е. Расстояние между двумя камерами). В этом случае стереоопределение становится неэффективным, и необходимо использовать монокулярные методы. На протяжении многих лет монокулярные и стереофонические VO почти прогрессировали как две независимые линии исследований. В оставшейся части этого раздела мы рассмотрели соответствующую работу в этих областях.

**Метод Стерео VO.**

Большая часть исследований, проведенных в VO, была произведена с использованием стереокамер. Основываясь на работе Моравека, Матти и Шафер [6], [7] использовали бинокулярную систему и процедуру Моравека обнаружения и отслеживания углов. Вместо использования скалярного представления неопределенности, как это сделал Моравек, они воспользовались матрицей ковариации ошибок триангулированных признаков и включили ее в шаг оценки движения. По сравнению с Моравеком, они продемонстрировали более высокие результаты в восстановлении траектории для планетарного вездехода с относительной ошибкой 2% на пути 5,5 м.

Олсон и др. [9], [13] позже расширили эту работу путем введения датчика абсолютной ориентации (например, компаса или всенаправленной камеры) и использования детектора угла Форстннера, который значительно быстрее вычисляется, чем оператор Моравека. Они показали, что использование оценки движения с помощью камеры приводит к ошибкам накопления с суперлинейным ростом пройденного расстояния, что привело к увеличению ошибок ориентации. И наоборот, когда включен датчик абсолютной ориентации, рост ошибок может быть уменьшен до линейной функции пройденного расстояния. Это привело к погрешности относительной позиции 1: 2% на 20-метровой траектории.

Лакроикс и другие [8] реализовал стереофонический подход VO для планетарных вездеходов, аналогичный с описанным ранее. Разница заключается в выборе ключевых точек. Вместо использования детектора Форстнера они использовали плотную стереосистему и затем выбирали ключевые точки кандидата, анализируя корреляционную функцию вокруг ее максимумов – подход, который позже использовался в работах [14], [15] и других работах. Этот выбор был основан на наблюдении, что существует сильная корреляция между формой кривой корреляции и стандартным отклонением функции глубины. Это наблюдение позднее было использовано Ченгом и другими [16], [17] в их окончательной реализации VO на борту марсоходов. Они улучшили более раннюю реализацию Олсона и др. [9], [13] в двух областях. Во-первых, после использования детектора угла Харриса они использовали кривизну корреляционной функции вокруг объекта, как предложено Лакроиксом и др., чтобы определить матрицу ковариации ошибок точки изображения. Во-вторых, как было предложено Нистером и др. [1] они использовали случайный выбор консенсуса RANSAC [18] на этапе оценки движения наименьших квадратов для исключения отклонений.

Другой подход к оценке движения и удалению отклонений для вездехода был предложен Миллелой и Зигварт [14]. Они использовали подход Ши-Томаси [19] для определения угла и похожи на Лакроикс, они сохраняли эти точки с высокой степенью уверенности в карту несоответствий стереоизображений. Затем оценка движения была решена сначала с использованием наименьших квадратов, как в методах ранее, а затем путем итерационного алгоритма ближайшей точки (ICP[[5]](#footnote-6)) [20] – алгоритм, популярный в трехмерной регистрации лазерных сканирований, для уточнения положения. Для обеспечения надежности в ICP была включена стадия удаления отклонений.

До сих пор упомянутые работы имеют общее значение, что трехмерные точки триангулированы для каждой стереопары, а относительное движение решается как проблема регистрации (выравнивания) 3-D-to-3 D. Совершенно другой подход, предложенный в 2004 году Нистером и др. [1]. Их статья известна не только для введения термина VO, но и для обеспечения первой долгосрочной реализации в режиме реального времени с надежной схемой отклонения отброса. Нистер и др. улучшили предыдущие реализации в нескольких областях. Во-первых, вопреки всем предыдущим работам, они не отслеживали функции среди кадров, а обнаруживали функции (углы Харриса) независимо во всех кадрах и допускали только совпадения между функциями. Это позволяет избежать смещения функции во время отслеживания на основе взаимной корреляции. Во-вторых, они не вычисляли относительное движение как проблему отслеживания 3-D-to-3-D-точки, а как 3-D-двумерную (2D) оценку положения камеры (эти методы описаны в разделе «Оценка движения»). Наконец, они включили в RANSAC исключение отклонений (outlier rejection) как шаг метода оценки движения. Другая схема оценки движения была введена Компортом и др. (Comport et al.) [21]. Вместо того, чтобы использовать 3-D-to-3-D-точечную регистрацию или 3-D-to-2-D методы оценки положения камеры, они полагались на квадрангообразный тензор (quadrifocal tensor WTF???), который позволяет вычислять движение из 2-D-to-D совпадение изображений без триангуляции трехмерных точек в любую из стереопар. Преимущество использования непосредственно необработанных двумерных точек вместо триангулированных трехмерных точек заключается в более точном вычислении движения.

**Метод Моно (Monocular) VO**

Отличие от стереофонической схемы заключается в том, что в монокулярном VO как относительное движение, так и трехмерная структура должны вычисляться из двухмерных данных о местоположении. Поскольку абсолютная шкала неизвестна, расстояние между первыми двумя позициями камеры обычно устанавливается равным единице. По мере поступления нового изображения относительный масштаб и представление камеры относительно первых двух кадров определяются с использованием либо знания трехмерной структуры, либо трифокального тензора [[6]](#footnote-7)[22].

Успешные результаты с одной камерой на большие расстояния (до нескольких километров) были получены в последнее десятилетие с использованием как перспективных, так и всенаправленных камер [23] - [29]. Связанные работы можно разделить на три категории: методы выделения особенностей (Feature-Based[[7]](#footnote-8)), методы на основе внешнего вида (выделение общего объекта) и гибридные методы. Feature-Based методы основаны на характерных и повторяемых чертах, которые отслеживаются по кадрам; методы на основе внешнего вида используют информацию об интенсивности всех пикселей в изображении или подобластях; и гибридные методы используют комбинацию из двух предыдущих.

В первой категории - работы авторов в [1], [24], [25], [27], [30] - [32]. Первый крупномасштабный VO реального времени с одной камерой был представлен Нистером и др. [1]. Они использовали RANSAC для удаления отклонений и оценки 3-D-2-D камеры, чтобы вычислить новое предстоящее положение камеры. Новизна их работы - использование пятиточечного минимального решателя [33] для вычисления гипотез движения в RANSAC. После этой статьи метод пяти точек RANSAC стал очень популярным в VO и использовалися в нескольких других работах [23], [25], [27]. Корке и др. [24] обеспечил подход к монокулярному VO на основе всенаправленных изображений из катадиоптрической (Catadioptric[[8]](#footnote-9)) камеры и оптического потока. Lhuillier [25] и Mouragnon и др. [30] представили подход, основанный на локальной настройке оконного пакета для восстановления как движения, так и трехмерной карты (это означает, что настройка пакета выполняется над окном последних m кадров). Опять же, они использовали пятиточечный метод RANSAC [33], чтобы удалить отклонения. Tardif и др. [27] представил подход к VO на машине в течение очень длительного промежутка времени (2,5 км) без узла корректировки. Вопреки предыдущей работе они отделили оценку вращения и смещения. Вращение оценивалось с использованием точек на бесконечности, а смещение из восстановленной трехмерной карты. Ошибочные соответствия были удалены с помощью пятиточечного метода RANSAC.

Среди подходов, основанных на выделении особенностей или гибридных, приближенными являются работы авторов в [26], [28] и [29]. Goecke и др. [26] использовал преобразование Фурье-Меллина для регистрации перспективных изображений плоскости земли, взятых из автомобиля. Милфорд и Уайет [28] представили метод извлечения приблизительной информации о скорости вращения и смещении из одной перспективной камеры [[9]](#footnote-10)(perspective camera), установленной на автомобиле, которая затем использовалась в схеме RatSLAM[[10]](#footnote-11) [34]. Они использовали отслеживание шаблонов в центре сцены. Основным недостатком подхода, основанного на выделении общего объекта является то, что он не являются надежным к окклюзии[[11]](#footnote-12)(occlusion). По этой причине, Scaramuzza и Siegwart [29] использовали метод выделения общих объектов с изображения, чтобы оценить вращение автомобиля, и особенности с наземной плоскости, чтобы оценить смещение и абсолютный масштаб. Функциональный подход также использовался для обнаружения сбоев метода, основанного на выделении общих объектов.

Все подходы, упомянутые ранее, разработаны для неограниченного движения в 6 степенях свободы (DoF – degree of free). Однако несколько работ VO были специально разработаны для транспортных ср. с ограничениями в движении. Преимуществом является уменьшение времени вычислений и улучшение точности движения. Например, Liang and Pears [35], Ke и Kanade [36], Wang и др. [37], и Guerrero и др. [38] использовал гомографию (Homography[[12]](#footnote-13)) для оценки движения на доминирующей наземной плоскости. Scaramuzza и др. [31], [39] ввели одноточечное удаление отклонений RANSAC, основанное на неголономных ограничениях (nonholonomic constraints[[13]](#footnote-14)) транспортного средства, чтобы ускорить оценку при движении до 400 Гц. В последующей работе они показали, что неголономные ограничения позволяют восстановить абсолютный масштаб от одной камеры всякий раз, когда автомобиль совершает поворот [40]. Вслед за этой неголономные ограничения транспортного ср. также использовались Pretto и др. [32] для улучшения отслеживания особенностей и Fraundorfer и др. [41] для настройки оконного узла (см. Следующий раздел).

**Уменьшение Дрифта** (Drift | Concept Drift[[14]](#footnote-15))

Поскольку VO работает путем постепенного вычисления пути камеры (положение за положением), ошибки, возникающие с каждым новым движением от кадра к кадру, накапливаются со временем. Это порождает дрифт оценочной траектории от реального пути. Для некоторых приложений крайне важно, чтобы дрифт был как можно меньше, что можно сделать с помощью локальной оптимизации по последним позициям камеры m. Этот подход, называется скользящей оконной связкой или регулировкой оконного узла, использовался в нескольких работах, таких как [41] - [44]. В частности, на 10-километровом эксперименте VO Konolige et al. [43] продемонстрировал, что регулировка оконного узла может уменьшить ошибку конечного положения в 2-5 раз. Очевидно, что дрифт VO также может быть уменьшен посредством комбинации с другими датчиками, такими как GPS и лазер, или даже с IMU [43], [45], [46].

**V-SLAM**[[15]](#footnote-16)

Хотя в этом учебном пособии основное внимание уделяется VO, стоит упомянуть параллельную линию исследований, проводимую визуальной одновременной локализацией и отображением (V-SLAM). Для углубленного изучения проблемы SLAM читатель ссылается на два учебника по этой теме Дурранта-Уайта и Бейли [47], [48]. В V-SLAM преобладают две методологии:

1) методы фильтрации объединяют информацию со всех изображений с распределением вероятности [49]

2) методы нефильтрования (также называемые ключевыми кадровыми методами) сохраняют оптимизацию глобальной привязки пакета к выбранным ключевым кадрам [50]. Основные преимущества обоих подходов были оценены и обобщены в [51].

В последние несколько лет успешные результаты были получены с использованием как одиночных, так и стереокамер [49], [52] - [62]. Большинство из этих работ были ограничены небольшими внутренними рабочими пространствами, и лишь некоторые из них недавно были разработаны для применения в широких областях [54], [60], [62]. Некоторые из ранних работ в V-SLAM реального времени были представлены Chiuso et al. [52], Динса [53] и Дэвисона [49] с использованием полноковариационного подхода Кальмана[[16]](#footnote-17). Преимущество работы Дэвисона состояло в том, чтобы объяснить повторяющуюся локализацию через произвольное время. Позже, Handa et al. [59] улучшили эту работу, используя метод активного согласования, основанный на вероятностной структуре. Civera et al. [60], основанный на этой работе, предложив комбинацию одноточечного RANSAC в калмановском фильтре, который использует имеющуюся ранее вероятностную информацию из фильтра на этапе гипотезы RANSAC. Наконец, Strasdat et al. [61] представил новую структуру для крупномасштабного V-SLAM, которая использует подход оптимизации ключевого кадра [50], принимая во внимание особый характер SLAM.

**VO в сравнении с V-SLAM**

В этом разделе анализируется связь VO с V-SLAM. Цель SLAM в целом (и, в частности, V-SLAM) - получить глобальную последовательную оценку пути робота. Это подразумевает сохранение следа карты окружающей среды (даже в том случае, когда карта сама по себе не нужна), потому что необходимо осознать, когда робот возвращается в ранее посещаемую область. (Это называется замыканием цикла. Когда обнаружено замыкание цикла, эта информация используется для уменьшения дрифта как по пути карты, так и по камере. Понимание того, когда происходит замыкание цикла и эффективное интегрирование этого нового ограничения в текущую карту, является двумя основными проблемами в SLAM.) И наоборот, VO стремится к постепенному восстановлению пути, позиция за позицией, и потенциальной оптимизации только по последним n позициям пути (это также называется установкой оконного узла). Оптимизацию этого скользящего окна можно считать эквивалентной построению локальной карты в SLAM; однако философия различна: в VO мы заботимся только о локальной согласованности траектории, а локальная карта используется для получения более точной оценки локальной траектории (например, при корректировке узла), тогда как SLAM связана с согласованностью глобальной карты.

VO может использоваться как строительный блок для полного алгоритма SLAM для восстановления инкрементного движения камеры; однако, чтобы сделать полный метод SLAM, нужно также добавить способ обнаружения замыкания цикла и, возможно, глобального шага оптимизации для получения метрически согласованной карты (без этого шага карта по-прежнему топологически согласована).

Если пользователю интересен только путь камеры и не интересна карта среды окружения, все же существует возможность использовать полный метод V-SLAM вместо одного из методов VO, описанных в этом учебном пособии. Метод V-SLAM потенциально намного более точен, потому что он налагает на путь гораздо больше ограничений, но не обязательно более надежный (например, отклонения при закрытии цикла могут серьезно повлиять на согласованность карты). Кроме того, он более сложный и вычислительно дорогой.

В конечном итоге выбор между VO и V-SLAM зависит от компромисса между производительностью и согласованностью и простотой реализации. Хотя глобальная согласованность пути камеры иногда желательна, VO торгует согласованностью для производительности в реальном времени, без необходимости отслеживать всю предыдущую историю камеры.

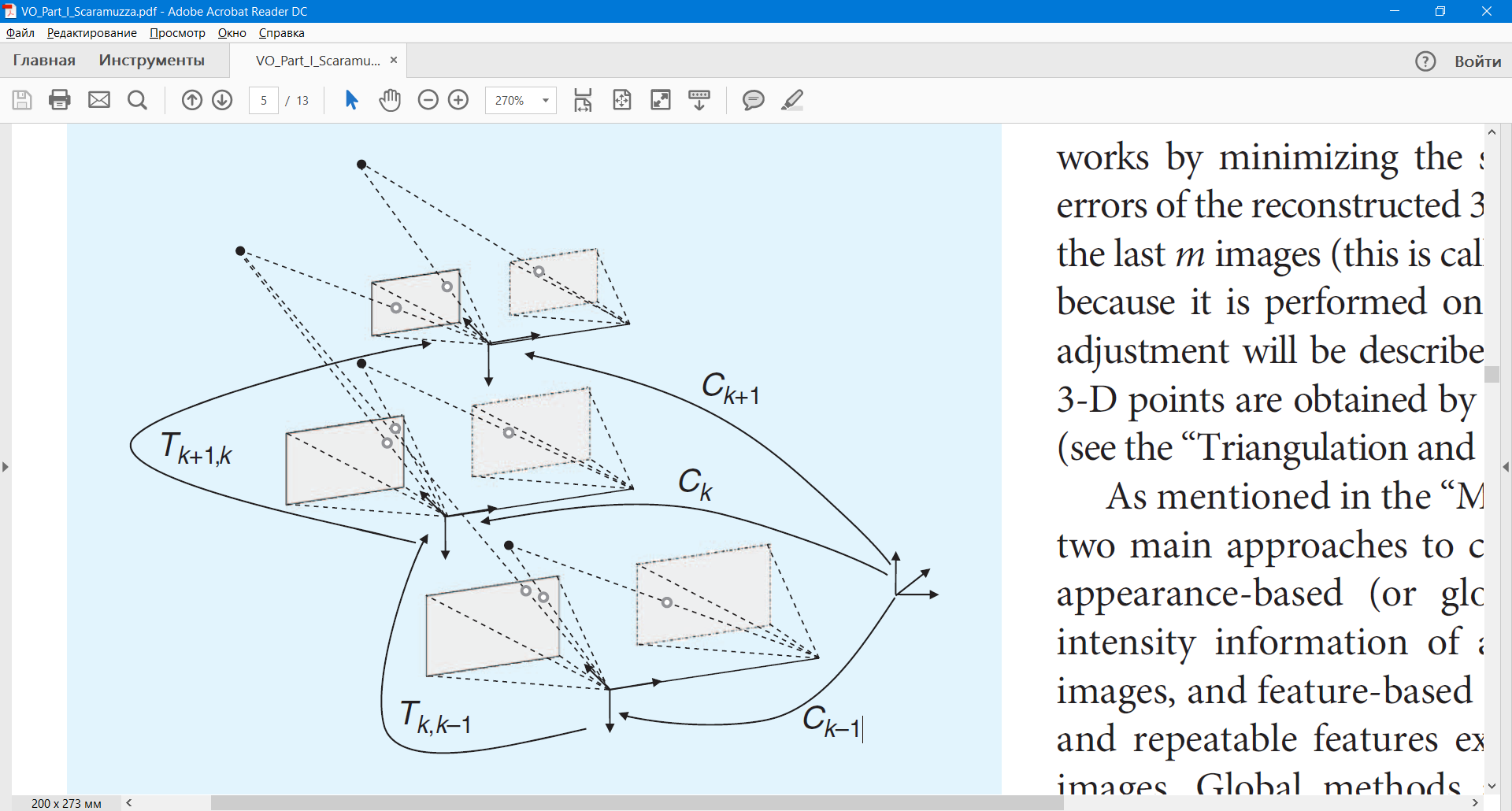


Рисунок 1. Иллюстрирует проблемы VO. Относительные позы Tk, k-1 смежных положений камеры (или положения системы камер) вычисляются по визуальным признакам и объединяются, чтобы получить абсолютные позы Ck относительно исходного координатного кадра при k = 0.

Формулировка проблемы VO.

Агент перемещается через среду и принимает изображения с жестко прикрепленной системой камеры в моменты времени k. В случае монокулярной системы набор изображений, взятых в моменты времени k, обозначается через . В случае стереосистемы в каждый момент времени есть левое и правое изображение, обозначаемое знаком и . На рисунке 1 показана иллюстрация этого параметра.

Для простоты предполагается, что кадр координат камеры также является рамкой координат агента(For simplicity, the camera coordinate frame is assumed to be also the agent’s coordinate frame).

В случае стереосистемы без ограничения общности система координат левой камеры может использоваться как источник.

Две позиции камеры при смежных моментах времени k – 1 и k связаны преобразованием закрепленного тела следующего вида: (1)

где (3) – матрица вращения и вектор трансляции. Множество содержит все последующие движения. Чтобы упростить обозначение, с этого момента будет использоваться вместо . Наконец, набор позиций камеры содержит преобразования камеры относительно начального координатного кадра при k = 0. Tекущая позиция может быть вычислена путем объединения всех преобразований и, следовательно, , причем является позицией камеры в момент k = 0, которое может быть задано пользователем произвольно.

Основной задачей в VO является вычисление относительных преобразований из изображений и , а затем для объединения преобразований для восстановления полной траектории камеры. Это означает, что VO постепенно восстанавливает путь, позиция за поизицией. После этого шага можно выполнить итерационное уточнение по последним m позициям, чтобы получить более точную оценку локальной траектории. Это итеративное уточнение работает, сводя к минимуму сумму квадратов ошибок реконкирования восстановленных трехмерных точек (т. Е. 3-Dmap) над последними m изображениями (это называется регулировкой оконного пучка, потому что оно выполняется в окне m frames. Настройка Bundle будет описана в части II этого урока). Трехмерные точки получают путем триангуляции точек изображения (см. Раздел «Выбор триангуляции и выбора ключевого кадра»).

Как упоминалось в разделе «Моно VO», существуют два основных подхода к вычислению относительных движений : основанные на внешнем виде – выделение общего объекта (или глобальные), которые используют информацию об интенсивности всех пикселей в двух входных изображениях, и методы основанные на выделении особенностей, которые используют только характерные и повторяемые функции, извлеченные (или отслеживаемые) через изображения. Глобальные методы менее точны, чем Feature-Based – методы выделения особенностей, и являются более дорогостоящими. (Как отмечалось в разделе «История VO», большинство методов, основанных на внешнем виде, были применены к монокулярному VO, что объясняется простотой реализации по сравнению со случаем стереофонической камеры.) Методы, основанные на выделении особенностей, требуют способности устойчиво сочетаться (или отслеживаться) в разных кадрах, но быстрее и точнее глобальных методов. Поэтому большинство реализаций VO основаны на Feature-Based методах.

Схема работы VO представлена на рисунке 2. Для каждого нового изображения (или пары изображений в случае стереокамеры) первые два шага состоят в обнаружении и сопоставлении двухмерных особенностей (характерные черты) с особенностями из предыдущих кадров. Двумерные особенности, являющиеся перепрограммированием одной и той же трехмерной особенности в разных кадрах, называются соответствиями изображений. (Как будет объяснено в части II этого урока, мы различаем соответствие особенностей и отслеживание функций. Первый состоит из обнаружения особенностей независимо во всех изображениях, а затем их сопоставления на основе некоторых показателей сходства, а второй - поиска особенностей на одном изображении, а затем отслеживать их на следующих изображениях с использованием метода локального поиска, такого как корреляция.) Третий шаг состоит в вычислении относительного движения между моментом времени k-1 и k. В зависимости от того, указаны ли соответствия в трех или двух измерениях, существует три различных подхода к решению этой проблемы (см. Раздел «Оценка движения»). Представление камеры затем вычисляется путем конкатенации с предыдущей позицией. Наконец, итеративное уточнение (корректировка пучка) может быть выполнено в течение последних m кадров для получения более точной оценки локальной траектории. Оценка движения объясняется в этом уроке (см. Раздел «Оценка движения»). Обнаружение и согласование функций и комплектация будут описаны в части II. Также обратите внимание на то, что для точного вычисления движения соответствующие соответствия функций не должны содержать выбросы (т. Е. Неправильные ассоциации данных). Обеспечение точной оценки движения при наличии выбросов является задачей надежной оценки, которая будет описана в части II этого урока. Большинство реализаций VO предполагают, что камера откалибрована. С этой целью в следующем разделе рассматриваются стандартные модели и процедуры калибровки для перспективных и всенаправленных камер.

Моделирование и калибровка камеры.

VO может быть выполнено с использованием как перспективных, так и всенаправленных камер. В этом разделе мы рассмотрим основные модели.

**Модель перспективной камеры**

Самая используемая модель для перспективной камеры предполагает систему прожекторов: изображение формируется путем пересечения лучей света от объектов через центр линзы (центр проекции) с фокальной плоскостью [Рисунок 3 (a)]. Пусть Точка сцены в камере отсчета и его проекция на плоскости изображения, измеренной в пикселях. Отображение из трехмерного мира на двумерное изображение дается с помощью уравнения проекции в перспективе: , (2)

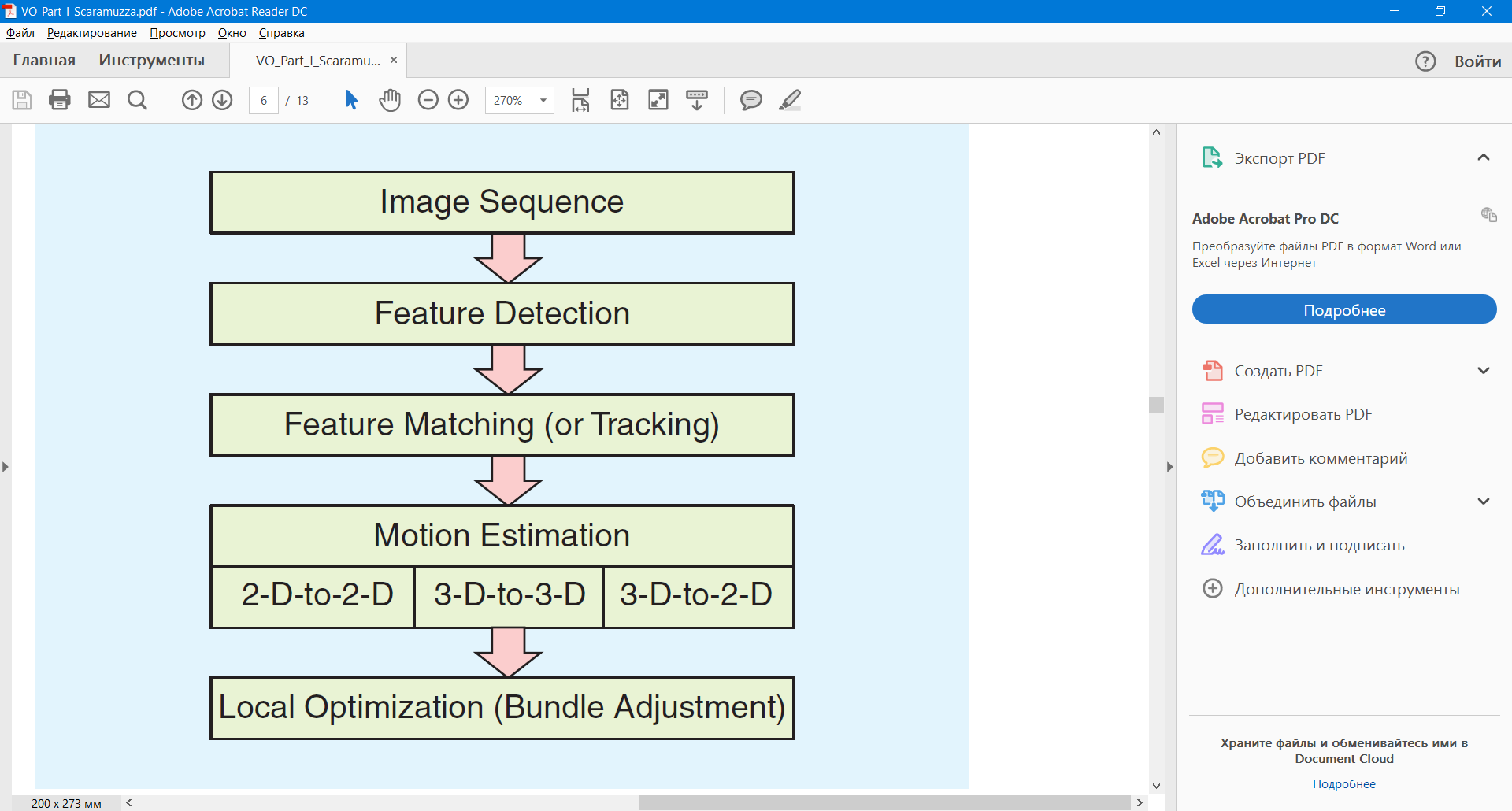
где - коэффициент глубины, и - фокусные расстояния, , – координаты изображения проекционного центра. Эти параметры называются внутренними параметрами. Когда поле зрения камеры больше 45 °, эффекты радиального искажения могут стать видимыми и могут быть смоделированы с использованием полинома второго или более высокого порядка. Вывод полной модели можно найти в учебниках компьютерного зрения, таких как [22] и [63]. Пусть - нормированные координаты изображения. Нормированные координаты будут использоваться в следующих разделах.

Рисунок 2.

Блок-схема, показывающая основные компоненты VO.

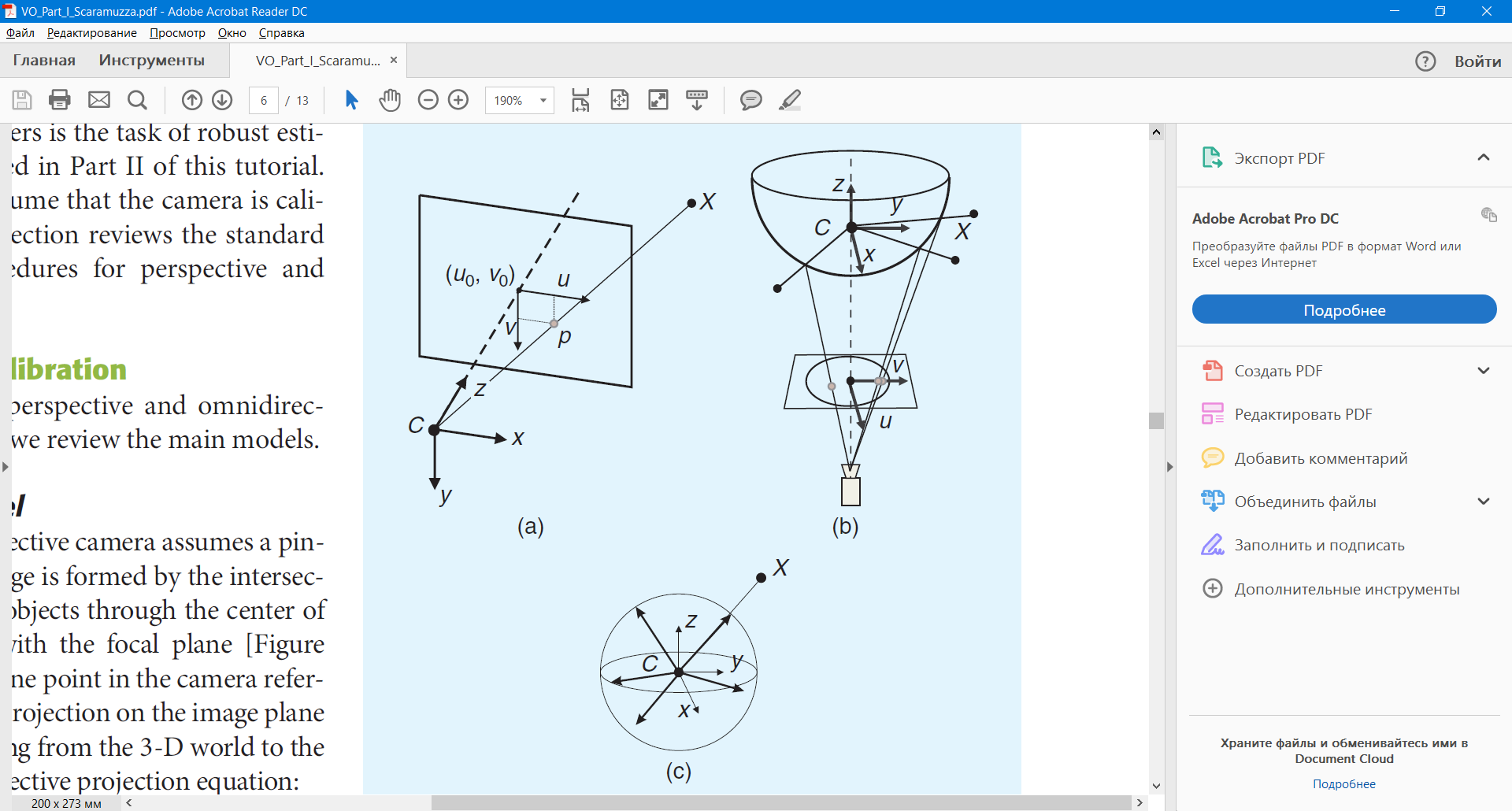


Рисунок 3. (a) Перспективная проекция, (b) катадиоптрическая проекция и (c) сферическая модель для перспективных и всенаправленных камер. Точки изображения представляются в виде направлений к просматриваемым точкам, нормированным на единичной сфере.

**Модель всенаправленной камеры**

Всенаправленные камеры - камеры с широким полем зрения (даже более 180) и могут быть построены с использованием оптических линз или сочетанием стандартных камер с зеркалами [последние называются катадиоптрическими камерами, рис. 3 (b)]. Типичные формы зеркала в катадоптических камерах представляют собой квадратичные поверхности вращения (например, параболоидные или гиперболоидные), поскольку они гарантируют единый центр проекции, что позволяет использовать теорию оценки движения, представленную в разделе «Оценка движения». В настоящее время существует две приемлемые модели для всенаправленных камер. Первый из них, предложенный Гейер и Даниилидис [64], предназначен для общих катадиоптрических камер (параболических или гиперболических), а второй - для Scaramuzza et al. [65] является унифицированной моделью как для глазных, так и для катадоптических камер. Обзор этих двух моделей можно найти в [66] и [67]. Уравнение проекции унифицированной модели выглядит следующим образом: , (3)

где и являются внутренними параметрами, которые зависят от типа зеркала или линзы типа рыбий глаз. Как показано в [65], n=4 является разумным выбором для большого разнообразия зеркал и линз для рыбьего глаза. Наконец, эта модель предполагает, что плоскость изображения удовлетворяет идеальному свойству, что оси симметрии камеры и зеркала выровнены. Несмотря на то, что это допущение справедливо для большинства катадиоптрических камер и камер с линзой рыбий глаз, несбалансированность можно моделировать, введя перспективную проекцию между идеальной и реальной плоскостью изображения [66].

**Сферическая модель**

Как упоминалось ранее, желательно, чтобы камера имела один проекционный центр (также называемый одной эффективной точкой обзора). В катадиоптрической камере это происходит, когда лучи, отраженные зеркалом, пересекают все в одной точке (а именно С). Существование этой точки позволяет нам моделировать любую всенаправленную проекцию как отображение с единственной точки зрения в сферу. Для удобства обычно применяется единичная сфера. Важно заметить, что сферическая модель применяется не только к всенаправленным камерам, но и к перспективным камерам. Если камера откалибрована, любая точка в перспективе или всенаправленном изображении может быть отображена в вектор на сфере единицы. Как можно видеть на рис. 3 (с), эти единичные векторы представляют собой направления к просматриваемым точкам сцены. Эти векторы называются нормализованными точками изображения на единичной сфере.

**Калибровка камеры**

Целью калибровки является точное измерение внутренних и внешних параметров системы камеры. В многокамерной системе (например, стерео и тринокуляр) внешние параметры описывают взаимное расположение и ориентацию между каждой парой камер. Самый популярный метод использует планарный шаблон, подобный шахматной доске. Известно положение квадратов на доске. Чтобы точно вычислить параметры калибровки, пользователь должен сделать несколько снимков доски, показанной в разных положениях и ориентации, гарантируя, что поле зрения камеры заполнено как можно больше. Внутренние и внешние параметры затем обнаруживаются с помощью метода минимизации наименьших квадратов. Входными данными являются 2-D положения углов квадратов доски и их соответствующие координаты пикселей в каждом изображении. Для MATLAB и C. было разработано множество инструментов для калибровки камер. Обновленный список можно найти в [68]. Среди них наиболее популярные для MATLAB приведены в [69] и [70] - [72] - для перспективных и всенаправленных камер, соответственно. C реализация калибровки камеры для перспективных камер можно найти в OpenCV [73], библиотеке видения компьютера с открытым исходным кодом.

Оценка движения

Оценка движения - это основной шаг вычисления, выполняемый для каждого изображения в системе VO. Точнее, на этапе оценки движения вычисляется движение камеры между текущим изображением и предыдущим изображением. Благодаря объединению всех этих движений полная траектория камеры и агента (при условии, что камера жестко закреплена) может быть восстановлена. В этом разделе объясняется, как преобразование между двумя изображениями и может быть вычислено из двух наборов соответствующих признаков 0 , в моменты времени k-1 и k соответственно. В зависимости от того, указаны ли соответствия особенностей в двух или трех измерениях, существует три разных метода.

* 2-D-to-2-D: В этом случае оба и указаны в двухмерных координатах изображения.
* 3-D-to-3-D: В этом случае оба и указаны в 3-D. Для этого необходимо триангулировать 3-D точки в каждый момент времени; например, с помощью системы стереокамер.
* 3-D-to-2-D: В этом случае задаются в 3-D, а - их соответствующие двумерные повторы на изображении . В монокулярном случае трехмерная структура должна быть триангулирована из двух соседних видов камеры (например, и ), а затем сопоставляется с функциями двумерного изображения в третьем представлении (например, ). В монокулярной схеме необходимы совпадения по меньшей мере трех точек зрения.

Обратите внимание, что особенности (особенные черты) могут быть точками или линиями. В общем, из-за отсутствия линий в неструктурированных сценах в VO используются точечные особенности. Подробный обзор этих трех подходов как для точечных, так и для линейных особенностей можно найти в [74]. Формулировка, приведенная в этом руководстве, предназначена только для точечных особенностей.

**2-D-2-D: Движение от изображения**

**Соответствие особенностей**

Оценка основной матрицы

Геометрические отношения между двумя изображениями и калиброванной камеры описываются так называемой существенной матрицей E. E содержит параметры движения камеры до неизвестного масштабного коэффициента для преобразования в следующую форму: , (4)

где , и

. (5)

Символ используется для обозначения того, что эквивалентность действительна до мультипликативного скаляра.

Существенную матрицу можно вычислить из соответствия двух-двухзначных признаков, и вращение, и преобразование можно непосредственно извлечь из E. Основным свойством оценки движения на основе 2-D-to-2-D является эпиполярное ограничение, которое определяет линию, на которой соответствующая точка объекта из лежит на другом изображении (рис. 4). Это ограничение можно сформулировать с помощью , где – местоположение объекта на одном изображении (например, ), а – местоположение его соответствующей особенности в другом изображении (например, ) , и - нормированные координаты изображения. Для простоты в следующих разделах будут использоваться нормированные координаты в виде (См. Раздел «Модель перспективной камеры»). Однако очень похожие уравнения могут быть также получены для нормированных координат на единичной сфере (см. Раздел «Сферическая модель»).

Существенную матрицу можно вычислить из соответствия двух-двухзначных признаков с использованием эпиполярного ограничения. Решение минимального случая включает в себя пять соотношений 2-D-to-2-D [75] и эффективную реализацию, предложенную Нистером в [76]. Пятиточечный алгоритм Нистера стал стандартом для оценки движения 2-D-to-2-D при наличии отклонений (проблема надежной оценки будет рассмотрена во второй части этого урока). Простое и прямое решение для некомпланарных точек - восьмиточечный алгоритм Лонгета-Хиггинса [2], который кратко излагается здесь. Каждое совпадение характеристик дает ограничение следующего вида:

, (6)

где .

Стэк ограничений из восьми точек дает систему линейных уравнений , и, решая систему, параметры E могут быть вычислены. Эта однородная система уравнений может быть легко решена с помощью разложения по сингулярным значениям (SVD) [2]. Наличие более восьми точек приводит к переопределенной системе для решения методом наименьших квадратов и обеспечивает степень надежности шума. SVD A имеет вид , а оценка наименьших квадратов E с может быть найдена как последний столбец V. Однако эта линейная оценка E делает не соответствуют внутренним ограничениям существенной матрицы, которые исходят из умножения матрицы вращения R и кососимметрической матрицы сдвига . Эти ограничения видны в сингулярных значениях существенной матрицы. Действительная существенная матрица после SVD равна и имеет , что означает, что первое и второе сингулярные значения равны, а третье равно нулю. Чтобы получить действительную E, удовлетворяющую ограничениям, решение нужно проецировать на пространство действительных существенных матриц. Проецируемая существенная матрица равна . Заметим, что решение восьмиточечного алгоритма вырождается, когда трехмерные точки копланарны. И наоборот, пятиточечный алгоритм работает и для копланарных точек. Наконец, обратите внимание, что восьмиточечный алгоритм работает как для откалиброванных (перспективных, так и всенаправленных) и некалиброванных (только перспективных) камер, тогда как пятиточечный алгоритм предполагает, что камера (перспективная или всенаправленная) откалибрована.

Извлечение R и t из E

Из оценки части вращения и преобразования можно извлечь. В общем, существует четыре различных решения для R, t для одной существенной матрицы; однако, путем триангуляции одной точки, можно определить правильную пару R, t. Четыре решения:

,

,

где , (7)

Описано эффективное разложение E в R и t в работе [76].

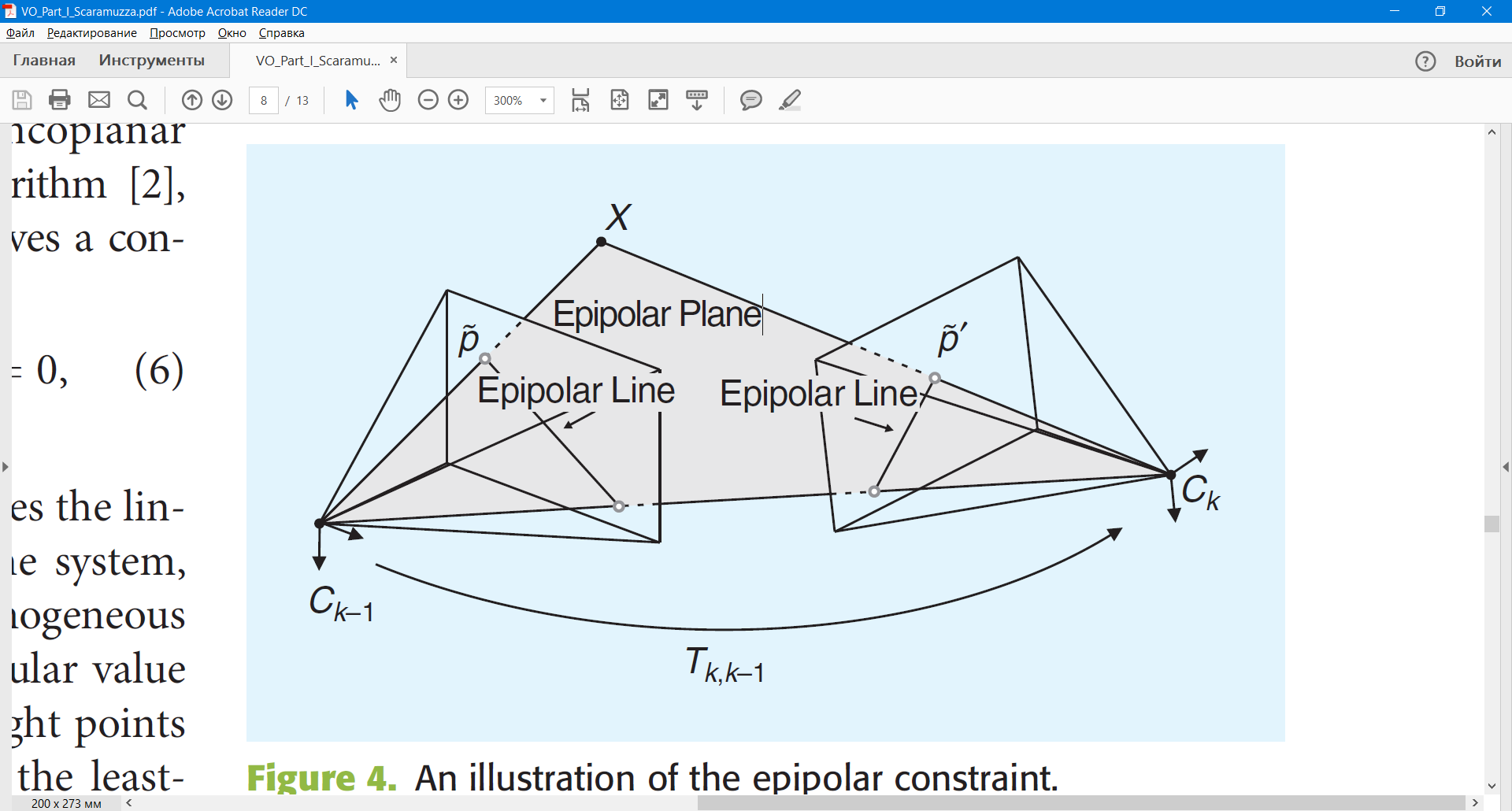


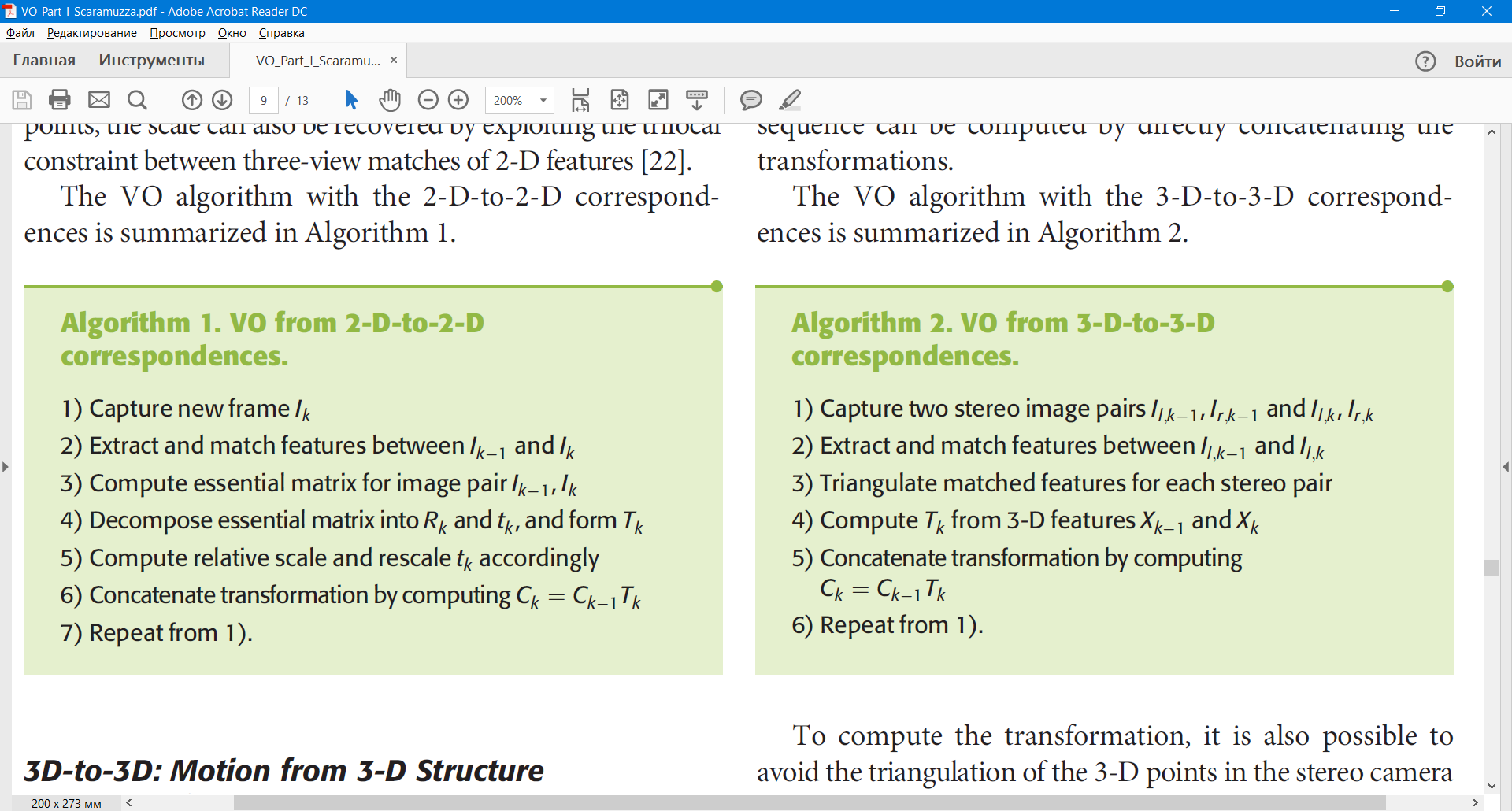
Рисунок 4. Изображение эпиполярного ограничения.

После выбора правильного решения путем триангуляции точки и выбора решения, где точка находится перед обеими камерами, нелинейная оптимизация параметров вращения и трансляции должна выполняться с использованием оценки R, t в качестве начальных значений. Функция для сведения к минимуму - это ошибка повторения, определенная в (10).

Вычисление относительной шкалы

Чтобы восстановить траекторию последовательности изображений, различные преобразования должны быть объединены. Для этого необходимо вычислить надлежащие относительные шкалы, поскольку абсолютный масштаб перевода не может быть вычислен из двух изображений. Тем не менее, можно вычислить относительные масштабы для последующих преобразований. Один из способов сделать это - триангулировать 3-D точки и из двух последующих пар изображений. Из соответствующих трехмерных точек можно вычислить относительные расстояния между любой комбинацией двух трехмерных точек. Соответствующую шкалу можно затем определить из отношения расстояния r между парой точек в и пары в .

, (8)



Для обеспечения надежности вычисляются коэффициенты масштабирования для многих пар точек и используется среднее значение (или в присутствии отклонеий, медиана). Затем вектор трансляции t масштабируется с этим отношением расстояний. Обратите внимание, что для вычисления относительных масштабов требуются особенности, которые должны быть сопоставлены (или отслеживаться) по нескольким кадрам (по крайней мере три). Вместо выполнения явной триангуляции трехмерных точек масштаб можно также восстановить, используя трифокальное ограничение между трехмерными совпадениями двухмерных особенностей [22].

Алгоритм 1 это обобщение алгоритмов VO и 2-D-to-2-D.

**3D-to-3D: движение от трехмерной структуры**

**Соответствия**

Для случая соответствующих 3-D-to-3-D признаков движение камеры Tk можно вычислить, определив выравнивающее преобразование двух трехмерных наборов соответствий. Соответствующие черты 3-D-to-3-D доступны при стереофоническом зрении.

Общее решение состоит в нахождении , который минимизирует расстояние между двумя трехмерными наборами признаков

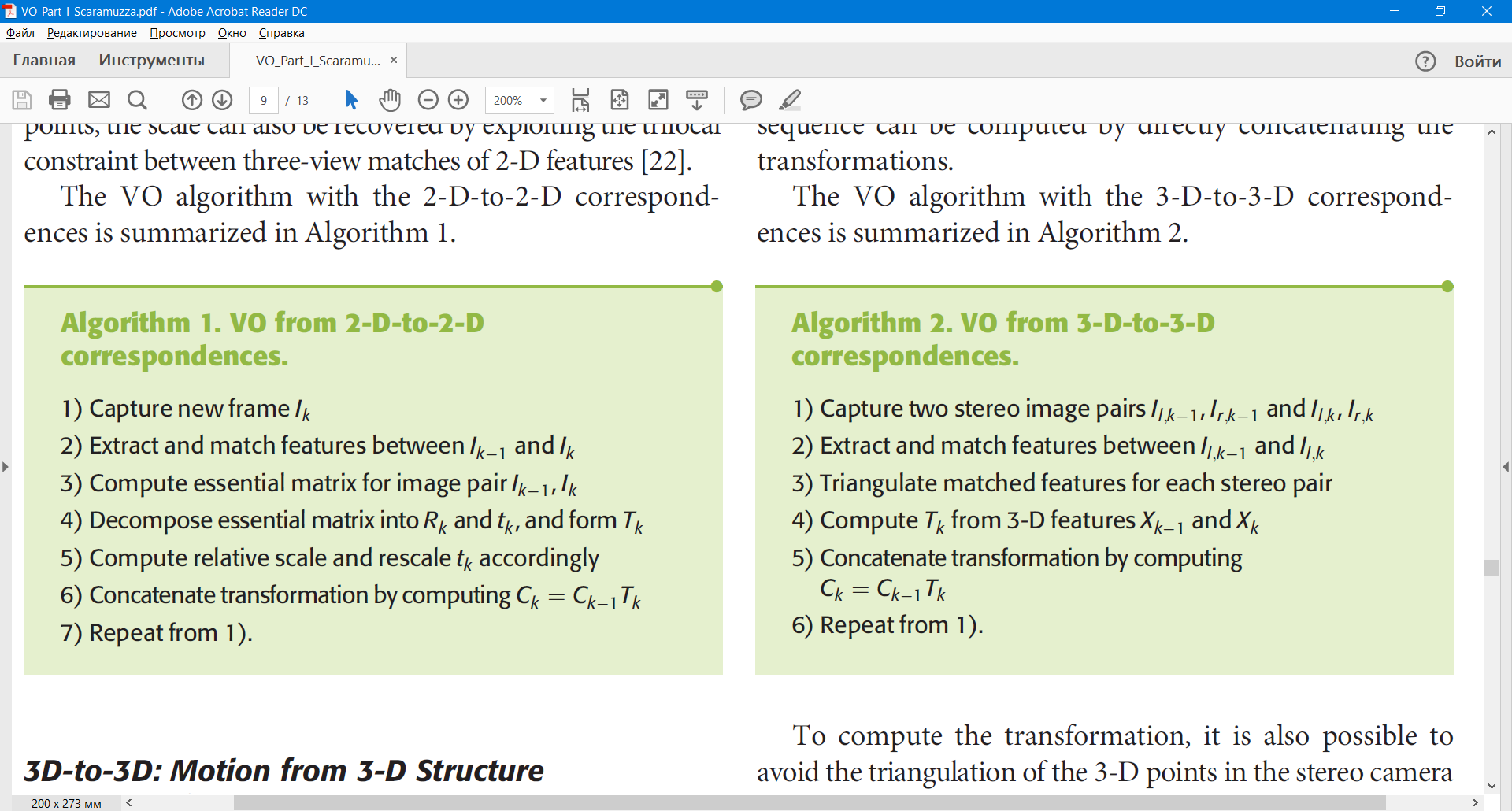
(9)

где верхний индекс i обозначает i-й признак совпадения, а и - однородные координаты трехмерных точек, т.е. .

Как показано в [77], решение минимального случая включает в себя три 3-D-to-3-D неколлинеарные соответствия, которые могут быть использованы для надежной оценки при наличии отклонений (часть II этого урока). Для случая соответствий, одно из возможных решений (согласно Arun и др. [78]) заключается в вычислении части преобразования как разности центроидов трехмерных наборов признаков и части вращения с использованием SVD. Преобразование задается , где означает среднее арифметическое значение.

Вращение может быть эффективно вычислено с использованием SVD как , где и и являются множествами соответствующих трехмерных точек.

Если известны неопределенности измерений трехмерных точек, их можно добавить в качестве весов в оценку, как описано Маймоном и др. [17]. Вычисленные преобразования имеют абсолютный масштаб, и, следовательно, траектория последовательности может быть вычислена путем непосредственной конкатенации преобразований.



В алгоритме 2 обобщен алгоритм VO с 3-D-to-3-D соответствием.

Чтобы вычислить преобразование, можно также избежать триангуляции трехмерных точек в стереокамере и вместо этого использовать квадрифокальные ограничения. Этот метод был отмечен Comport et al. [21]. Квадрифокальный тензор позволяет вычислить преобразование непосредственно из 2-D-to-2-D стереосоответствий.

**3-D-to-2-D: Движение от трехмерной структуры и соответствия функций изображения**

Как указывал Nister et al. [1], оценка движения от 3-D-to-2-D соответствий является более точной, чем от 3-D-to-3-D соответствий, поскольку она минимизирует ошибку репроекции[[17]](#footnote-18) изображения (10) вместо 3-D-to -3-D ошибка положения особенности (9). Преобразование вычисляется из 3-D-to-2-D соответствий и : может быть оценено по стереоданным данным или, в монокулярном случае, из триангуляции измерений изображения и . Последнее, однако, требует соответствия изображений по трем представлениям.

Общая формулировка в этом случае состоит в том, чтобы найти , который минимизирует ошибку воспроизведения изображения

, (10)

- где перепроекция трехмерной точки в изображение согласно преобразованию . Эта проблема известна как перспектива из n точек (PnP) (или резекции), и в литературе существует множество различных решений [79]. Как показано в [18], минимальный случай включает в себя три 3-D-to-2-D соответствия. Это называется перспективой из трех точек (P3P) и возвращает четыре решения, которые можно устранить с помощью одной или нескольких дополнительных точек. (Быстрая реализация P3P описана в [80], а код C можно свободно загрузить с веб-страницы авторов.) В случае 3-D-to-2-D P3P является стандартным методом для надежной оценки движения при наличии выбросов [18]. Надежная оценка будет описана в части II этого урока. Простое и прямое решение задачи PnP для точек - алгоритм прямого линейного преобразования [22]. Одно соответствие 3-D-to-D-точка обеспечивает два ограничения следующего вида для записей .

, (11)

где каждый представляет собой четыре вектора (j-я строка ) и x, y, z - координаты 3-D точки .

Стек ограничений шеститочечных соответствий дает линейную систему уравнений вида . Входы P могут быть вычислены из нулевого вектора A, например, с использованием SVD. Части вращения и трансляции можно легко извлечь из . Полученное вращение R не обязательно ортонормировано. Однако это не проблема, так как и R, и t могут быть уточнены нелинейной оптимизацией ошибки повторения, как определено в (10).

Оценка движения 3-D-to-2-D предполагает, что двухмерные точки изображения поступают только от одной камеры. Это означает, что для случая стереокамеры, 2-D точки изображения - это как левая, так и правая камера. Очевидно, что желательно одновременно использовать точки изображения обеих камер. Обобщенная версия алгоритма оценки движения 3-D-2-D для неконтактных лучей (т. е. двухмерных точек изображения из нескольких камер) была предложена Нистером в [81] для камер, калиброванных по внешнему виду (т.е. взаимная позиция и ориентация между камерами известна).

Для монокулярного случая необходимо триангулировать 3-мерные точки и поочередно оценивать положение от 3-D-to-2-D совпадений. Эта чередующаяся схема часто упоминается как SFM. Начиная с двух видов, исходный набор трехмерных точек и первое преобразование вычисляются из совпадений особенностей 2-D-to-2-D.

Последующие преобразования затем вычисляются из совпадений особенностей 3-D-to-2-D. Для этого особенности (черты) должны быть сопоставлены (или отслеживаться) по нескольким кадрам (по крайней мере три). Новые трехмерные особенности снова триангулируются, когда новое преобразование вычисляется и добавляется к набору трехмерных особенностей. Основная задача этого метода состоит в том, чтобы поддерживать согласованный и точный набор триангулированных трехмерных функций и создавать совпадения функций 3-D-to-2-D для по меньшей мере трех смежных кадров.

Алгоритм 3 алгоритма VO с 3-D-to-2-D соответствиями обобщен.

Выбор триангуляции и ключевого кадра

Некоторые из предыдущих методов оценки движения требуют

Выбор триангуляции и ключевого кадра

Некоторые из предыдущих методов оценки движения требуют триангуляции трехмерных точек (структуры) из двухмерных изображений. Для вычисления более точной оценки локальной траектории требуется также вычисление компоновки путем корректировки пучка (часть II этого урока).

Триангулированные 3-D точки определяются пересечением обратных проецируемых лучей из двухмерных изображений, соответствующих по меньшей мере двум кадрам изображения. В идеальных условиях эти лучи пересекались бы в одной трехмерной точке. Однако из-за шумов изображения, моделей камеры и ошибок калибровки, а также неопределенности соответствия, они никогда не пересекаются. Следовательно, точка на минимальном расстоянии, в смысле наименьших квадратов, от всех пересекающихся лучей может быть взята как оценка трехмерного точечного положения. Обратите внимание, что стандартное отклонение расстояний трехгранной трехмерной точки от всех лучей дает представление о качестве трехмерной точки. Трехмерные точки с большой неопределенностью будут выброшены. Это происходит, особенно когда кадры принимаются с очень близкими интервалами по сравнению с расстоянием до точек сцены. Когда это происходит, трехмерные точки демонстрируют очень большую неопределенность. Один из способов избежать этого состоит в том, чтобы пропускать кадры, пока средняя неопределенность трехмерных точек не будет ниже определенного порога. Выбранные кадры называются ключевыми кадрами. Выбор ключевого кадра - очень важный шаг в VO и всегда должен быть выполнен до обновления движения.

обсуждение

Согласно Nister et al. [1], есть преимущество в использовании методов 2-D-to-2-D и 3-D-2-D по сравнению с методом 3-D-to-3-D для вычисления движения. Нистер сравнил производительность VO 3-D-to-3-D с корпусом 3-D-to-2-D для системы стереокамер и обнаружил, что последнее значительно превосходит первое. Причина связана с тем, что триангулированные 3-D точки гораздо более неопределенны в направлении глубины. Когда в вычислении движения используются соответствия 3-D-to-3-D, их неопределенность может оказать разрушительное влияние на оценку движения. Фактически, в случае 3-D-to-3-D ошибка 3-D положения (9) минимизируется, тогда как в 3-D-to-2-D случае ошибка воспроизведения изображения, (10) ,

В монокулярной схеме метод 2-D-to-2-D является предпочтительным по сравнению со случаем 3-D-to-2-D, поскольку он избегает точечной триангуляции. Однако на практике метод 3-D-2-D используется чаще, чем метод 2-D-to-2-D. Причина кроется в его более быстрой ассоциации данных. Как будет описано в части II этого руководства, для точного вычисления движения крайне важно, чтобы входные данные не содержали выбросы. Отклонение выброса - очень деликатный шаг, а время вычисления этой операции

строго связано с минимальным количеством точек, необходимых для оценки движения. Как упоминалось ранее, для случая 2-D-to-2-D требуется минимум 5-точечных соответствий (см. Алгоритм с пятью точками); однако в случае 3-D-2-D движения (см. P3P) необходимы только три соответствия. Как будет показано в части II этого урока, это меньшее количество точек приводит к значительно более быстрой оценке движения.

Преимущество схемы стереокамер по сравнению с монокулярным, помимо того, что трехмерные функции вычисляются непосредственно в абсолютном масштабе, заключается в том, что совпадения должны вычисляться только между двумя представлениями вместо трех представлений, как в монокулярной схеме. Кроме того, поскольку трехмерная структура вычисляется непосредственно из одной стереопары, а не из соседних кадров, как в монокулярном случае, стерео-схема имеет меньший дрейф, чем монокулярный, в случае небольших движений. = Монокулярные методы интересны тем, что стерео VO вырождается в themonocular случае, когда расстояние до сцены намного больше, чем стерео базовая линия (т. Е. Расстояние между двумя камерами). В этом случае стереовидение становится неэффективным и монокулярным.

Независимо от выбранного метода вычисления движения, всегда необходимо выполнить локальную настройку расслоения (по последним m кадрам) для вычисления более точной оценки траектории. После корректировки связки эффекты метода оценки движения значительно облегчаются.

Выводы

В этом учебнике описана история ВО, постановка задачи и различные подходы к вычислению движения. VO - это хорошо понятая и установленная часть робототехники. В части II этого руководства будут обобщены оставшиеся строительные блоки конвейера VO: как выявлять и сопоставлять характерные и повторяемые функции в разных кадрах, надежную оценку при наличии выбросов и корректировку пакета. Кроме того, будут включены распространение ошибок, приложения и ссылки на бесплатный код для загрузки.

Подтверждения

Авторы выражают благодарность Константинос Дерпанис, Олегу Народицкому, Каролине Баэзу и Андреа Ценси за их плодотворные комментарии и предложения.

1. В **монокулярном VO** вы можете только сказать, что вы переместили одну единицу по x, две единицы по y и т. д. В то время как в **стерео**, вы можете сказать, что вы переместили один метр в x, два метра в y. [ [github](https://avisingh599.github.io/vision/visual-odometry-full/) ] [↑](#footnote-ref-2)
2. **Outline-Rejection** – Алгоритм для нахождения отклонения, которое очень сильно отличается от предыдущий наблюдений. [ [tutorial](https://archive.siam.org/meetings/sdm10/tutorial3.pdf) ] [↑](#footnote-ref-3)
3. ***RANdom SAmple Consensus*** *–* стабильный метод оценки параметров модели на основе [случайных выборок](https://ru.wikipedia.org/wiki/Выборка). Схема RANSAC устойчива к зашумлённости исходных данных. Метод был предложен в 1981 году Фишлером и Боллесом. [↑](#footnote-ref-4)
4. **SFM** – Structure from Motion [[sfm](https://en.wikipedia.org/wiki/Structure_from_motion)] [↑](#footnote-ref-5)
5. **Iterative closest point** ( **ICP** ) – алгоритм, используемый для [минимизации разницы между двумя облаками точек](https://en.wikipedia.org/wiki/Point_set_registration) . ICP часто используется для реконструирования 2D или 3D поверхностей из разных сканирований, локализации роботов и оптимального планирования пути (особенно, когда одометрия колес ненадежна из-за скользкой местности) и т.д. [↑](#footnote-ref-6)
6. Тензор – [ref](https://habr.com/post/261563/). [↑](#footnote-ref-7)
7. **The Feature-Based Appr oach** [[ref]](https://users.fmrib.ox.ac.uk/~steve/review/review/node2.html) [↑](#footnote-ref-8)
8. **катадиоптрические системы** — это разновидность оптических систем, содержащих в качестве оптических элементов как сферические зеркала(катоптрику), так и линзы. [↑](#footnote-ref-9)
9. [**perspective camera**](http://www1.cs.columbia.edu/~jebara/htmlpapers/SFM/node6.html) [↑](#footnote-ref-10)
10. **simultaneous localization and mapping** (**SLAM**) – это вычислительная проблема построения или обновления карты неизвестной среды, одновременно отслеживая местоположения внутри нее. RatSLAM - это реализация модели гиппокампа, которая может выполнять SLAM в реальном времени на реальном роботе. [↑](#footnote-ref-11)
11. **Occlusion** – случай, когда невозможно отследить какой либо предмет ели он находится за каким-либо существенным препятствием. [[ref]](https://stackoverflow.com/questions/2764238/image-processing-what-are-occlusions) [↑](#footnote-ref-12)
12. **Homography** – **Проективное преобразование** [проективной плоскости](https://ru.wikipedia.org/wiki/Проективная_плоскость) — это преобразование, переводящее прямые в прямые. (Изменение перспективы) [↑](#footnote-ref-13)
13. nonholonomic – [[ref]](https://ru.wikipedia.org/wiki/Неголономная_система). [↑](#footnote-ref-14)
14. **Concept Drift** означает, что статистические свойства целевой переменной, которую модель пытается предсказать, изменяются с течением времени непредвиденным образом. [[ref]](https://en.wikipedia.org/wiki/Concept_drift) [↑](#footnote-ref-15)
15. **SLAM** (см 9 сноску) [↑](#footnote-ref-16)
16. **Kalman filter** – [[ref]](https://www.bzarg.com/p/how-a-kalman-filter-works-in-pictures/). [↑](#footnote-ref-17)
17. Reprojection error – [[ref]](https://en.wikipedia.org/wiki/Reprojection_error) [↑](#footnote-ref-18)